Discrete Mathematics 26 (1979) 111–118

FAMILIES OF FINITE SETS SATISFYING A UNION CONDITION

Peter FRANKL

Mathematical Institute of the Hungarian Academy of Science, Budapest V, Realtanoda u. 13–15, Hungary

Received 10 June 1975
Revised 17 November 1978

Let n, t, k be integers, $n \geq t \geq 1, k \geq 2$. Let $x = \{1, 2, \ldots, n\}$. Let $\mathcal{F}$ be a family of subsets of x such that the cardinality of the union of any k members of $\mathcal{F}$ is at most $n-t$. How large $|\mathcal{F}|$ can be and which are the optimal families? We answer these questions for $t \leq 2^k k/150$.

1. Introduction

For i, j integers, let $[i, j]$ denote the set of integers in $i \leq h \leq j$. Let $\mathcal{F}$ be a family of subsets of $x = [1, n]$. If t, k are positive integers $k \geq 2, t \leq n$, then $\mathcal{F}$ is said to have property $P(n, k, t)$ if for $F_1, \ldots, F_k \in \mathcal{F}$ we always have $|\bigcup_{i=1}^k F_i| \leq n-t$. We say a family $\mathcal{F}$ is $E(n, k, t, s)$ (s is a non-negative integer) to mean that it has the form $\mathcal{F} = \{F \subseteq X \mid |F \cap Y| \leq s\}$ where $Y \subseteq X, |Y| = t + ks \leq n$. We denote by $e(n, k, t, s)$ the common cardinality of the $E(n, k, t, s)$ families. Clearly if $\mathcal{F} \in E(n, k, t, s)$, then $\mathcal{F}$ has $P(n, k, t)$ and $e(n, k, t, 0) = 2^{n-t}$.

P. Erdös and the author have the following:

Conjecture. *If n, k, t are given $n \geq t \geq 1, k \geq 2$ and $\mathcal{F}$ is a family of subsets of $x = [1, n]$ which has property $P(n, k, t)$ and which is of maximal cardinality, then there exists a non-negative integer s such that $\mathcal{F} \in E(n, k, t, s)$ unless $t = 1, k = 2$.*

In the case $k = 2$ the validity of the conjecture was proved by Katona [1]. The case $t = 1$ is trivial (cf. Erdös et al. [2, p. 319 (ii)]). The aim of this paper is to prove the following:

Theorem 2. *The conjecture holds for n, k, t whenever $k > 2, n \geq t$, and $t \leq k2^k/150$.*

Let $[x]([x]^*)$ denote the greatest (smallest) integer less (more) than or equal to x, resp.

2. Preliminary results

The following result was proved by Brace and Daykin [3].

Theorem. *Suppose that $\mathcal{F}$ is a family of subsets of $x = [1, n]$ and that $\bigcup_{F \in \mathcal{F}} F = X$.*

If $\mathcal{F}$ *has* $P(n, k, 1)$, *then*

$$|\mathcal{F}| \leq e(n, k, 1, 1) = (k+2)2^{n-k-1}. \tag{2}$$

Now we define an operation which was first by Kleitman [4]. Let $\mathcal{H}$ be a family of subsets of x. Suppose that i, j are integers, $1 \leq i < j \leq n$. Let us set $A_{i,j}(\mathcal{H}) = \{A_{i,j}(H) \mid H \in \mathcal{H}\}$ where

$$A_{i,j}(H) = \begin{cases} (H - \{j\}) \cup \{i\} & \text{if } \ j \in H, i \notin H, ((H - \{j\}) \cup \{i\}) \notin \mathcal{H}, \\ H & \text{otherwise.} \end{cases}$$

The following two propositions are easily verified.

Proposition 1. *If* $\mathcal{H}$ *has property* $P(n, k, t)$, *then* $A_{i,j}(\mathcal{H})$ *has* $P(n, k, t)$ *as well.*

Proposition 2. *If* $\mathcal{H}$ *has* $P(n, k, t)$ *and* $A_{i,j}(\mathcal{H})$ *is* $E(n, k, t, s)$ *for some* $s \geq 0$, *then* $\mathcal{H}$ *is* $E(n, k, t, s)$ *for the same* s.

Starting with a family $\mathcal{G}$ of subsets of x, having the property $P(n, k, t)$ and applying the operation $A_{i,j}$ repeatedly for all the pairs i, j $(1 \leq i < j \leq n)$ after a finite number of steps we obtain a family $\mathcal{F}$ which still has the property $P(n, k, t)$ and satisfies $A_{i,j}(\mathcal{F}) = \mathcal{F}$ for any i, j $(1 \leq i < j \leq n)$.

3. An inequality and some consequences

Let $\mathcal{F}$ be a family of subsets of x having $P(n, k, t)$, $k \geq 3$. Suppose that $|\mathcal{F}|$ is maximal. According to Section 2, we may assume that $A_{i,j}(\mathcal{F}) = \mathcal{F}$ for any $1 \leq i < j \leq n$. The maximality of $|\mathcal{F}|$ implies that $\mathcal{F}$ is a hereditary family of sets i.e. whenever $F \in \mathcal{F}$, $G \subseteq F$ we have $G \in \mathcal{F}$. Combining these properties we prove:

Proposition 3. *If* $\{i_1, \ldots, i_q\} = F \in \mathcal{F}$, $i_1 < i_2 < \cdots < i_q$, $G = \{j_1, \ldots, j_r\}$, $j_1 < j_2 < \cdots < j_r$, $r \leq q$ *and* $i_p \geq j_p$ *for* $p = 1, \ldots, r$, *then* $G \in \mathcal{F}$.

Proof. As $\mathcal{F}$ is hereditary so is $F' = \{i_1, \ldots, i_r\} \in \mathcal{F}$. Now use $A_{i,j}(\mathcal{F}) = \mathcal{F}$ for $(i, j) = (j_p, i_p)(p = 1, \ldots, r)$ and the statement follows. Let us set

$$b = \left[\frac{n-t}{k}\right].$$

Proposition 4. $F_1 = \{n - t - bk + 1, n - t - (b-1)k + 1, \ldots, n - t + 1\} \notin \mathcal{F}$.

Proof. If $F_1 \in \mathcal{F}$, then in view of Proposition 3 $F_i = ([1, n] \cap \{n - t - bk - (i-2), \ldots, n - t - (i-2)\}) \in \mathcal{F}$ for $i = 2, \ldots, k$ but $F_1 \cup F_2 \cup \cdots \cup F_k = [1, n - t + 1]$ contradicting the property $P(n, k, t)$.

Proposition 5. *For every $F \in \mathcal{F}$ there exists an integer $s, 0 \leqslant s \leqslant b$ such that*

$$|F \cap [n-t-ks+1, n]| \leqslant s. \tag{3}$$

Proof. Let $F = \{i_1, \ldots, i_q\}, i_1 < \cdots < i_q$. All we have to show is that there exists an integer $s, 0 \leqslant s \leqslant b$, such that $i_{q-s} \leqslant n-t-ks$ or that $q \leqslant b$. But the contrary means that $q \geqslant b+1$ and for $p = 0, 1, \ldots, b$ we have $i_{q-p} \geqslant n-t-pk+1$ implying $F \in \mathcal{F}$ which is a contradiction to Proposition 4.

Corollary 1. *For every $F \in \mathcal{F}$ there exists a non-negative integer $s, 0 \leqslant s \leqslant b$, such that*

$$|F \cap [n-t-ks+1, n]| = s. \tag{4}$$

Proof. Choose the smallest s for which (3) is fulfilled.

Let us count the number N_s of subsets of x for which (4) holds with a fixed s. As any subset F of X is uniquely determined by its intersections with $[1, n-t-ks]$ and $[n-t-ks+1, n]$ so we obtain

$$N_s = \binom{t+ks}{s} 2^{n-t-ks}. \tag{5}$$

Using eq. (5) and Corollary 1 we deduce

$$|\mathcal{F}| \leqslant N_0 + N, \quad \text{where} \quad N = N_1 + \cdots + N_b. \tag{6}$$

Let us examine the ratio of consecutive terms in N. If $t \geqslant k$, then

$$\frac{N_s}{N_{s+1}} = \frac{\binom{t+ks}{s} 2^{-ks}}{\binom{t+k(s+1)}{s+1} 2^{-k(s+1)}} = \frac{2^k(s+1)}{t+k(s+1)} \prod_{i=1}^{s} \frac{t+(k-1)s+i}{t+(k-1)(s+1)+i}$$

$$\geqslant \frac{2^k}{t}\left(\frac{s+1}{s+2}\right)^{s+1} > \frac{2^k}{te} = \rho. \tag{7}$$

Hence for $t < 2^k/e$ which means $\rho > 1$ we conclude from (6) that

$$|\mathcal{E}| < N_0 + (\rho-1)\rho N_1 = 2^{n-t}(1+\tau) \tag{8}$$

where $\tau = (t+k)/(2^k - te)$. Thus

$$|\mathcal{F}| < 2^{n-t+1} \quad \text{for} \quad \tau \leqslant 1. \tag{9}$$

Using this inequality we derive the following

Theorem 1. *Let $\mathcal{F}$ be a family of subsets of x. Suppose that $\mathcal{F}$ has property $P(n, k, t)$ and that*

$$k \geqslant 6, \qquad t \leqslant \frac{2^{k-1}-k+1}{e+1} - 1. \tag{10}$$

Then $|\mathcal{F}| \leqslant 2^{n-t}$ with equality holding if and only if $\mathcal{F} \in E(n, k, t, 0)$.

Proof. If $\mathcal{F}$ has property $P(n, k-1, t+1)$ and $t \geqslant k$, then in view of (10) we obtain applying (9) for the triple $(n, k-1, t+1)$ $|\mathcal{F}| < 2^{n-t}$. If $\mathcal{F}$ has not $P(n, k-1, t+1)$, then we can find sets $F_1, \ldots, F_{k-1}$ belonging to $\mathcal{F}$ such that their union is of cardinality $n-(t+1)+1) = n-t$. Let us set

$$Y = X - \bigcup_{i=1}^{k-1} F_i.$$

Then the property $P(n, k, t)$ implies that $F \cap Y = \emptyset$ for any $F \in \mathcal{F}$ and the assertion follows. As for $k \geqslant 6$ the inequality (10) is satisfied for $t = k$ so it suffices to prove that if the assertion of the theorem is true for the triple (n, k, t) for every $n \geqslant t$ and $t' < t$, then the assertion holds for the triple (n', k, t') whenever $n' \geqslant t'$.

In order to prove this let $\mathcal{F}$ be a family of subsets of $X' = [1, n']$ having the property $P(n', k, t')$. Then $\mathcal{F}$ can also be regarded as a family of subsets of $X = [1, n]$ where $n = n' - (t - t')$, and $\mathcal{F}$ has $P(n, k, t)$ whence either $|\mathcal{F}| < 2^{n-t} = 2^{n'-t'}$ or $\mathcal{F}$ is $E(n, k, t, 0)$ i.e. there exists a subset Y of X such that $|Y| = t$ and $\mathcal{F} = \{F \subseteq X \mid F \cap Y = \emptyset\}$. In this case we set $Y' = Y - [n-(t-t')+1, n]$ and obtain $\mathcal{F} = \{F \subseteq X' \mid F \cap Y' = \emptyset\}$.

4. A lemma

Let us set $c = t/2^k$. As for $k \leqslant 19$ we have

$$\frac{k \cdot 2^k}{150} < \frac{2^{k-1} - k + 1}{e+1} - 1$$

so for $6 \leqslant k \leqslant 19$ the statement of Theorem 2 follows from Theorem 1. As for $k \leqslant 5$ we have $t \leqslant 1$ we may assume $k \geqslant 20$ and $t \geqslant (2^{k+1} - k + 1)/(e+1)$, we use these facts without referring to them.

Lemma. *Let $\mathcal{F}$ be a family of subsets of x, having property $P(n, k, t)$ and suppose that $|\mathcal{F}|$ is maximal. Then for $p = [\log[12c]^* e + 2c \log e]^*$ it has not property $P(n, k-1, t+p)$.*

Proof. If $n < t+p$, then we have nothing to prove. So assume $n \geqslant t+p$. Suppose that $\mathcal{F}$ has $P(n, k-1, t+p)$. In this case we may apply the inequality (7) for the pair $(k-1, t+p)$ and obtain

$$\frac{|\mathcal{F}|}{2^n} \leqslant \sum_{s=0}^{h} \frac{\binom{t+p+(k-1)s}{s}}{2^{t+p+(k-1)s}}. \tag{11}$$

Let us set $\binom{t+p+(k-1)s}{s}2^{-t-p-(k-1)s}=d_s$. Then

$$d_s/d_{s+1}=\frac{(s+1)2^{k-1}}{t+p+(k-1)(s+1)}\prod_{i=1}^{s}\frac{t+p+(k-2)s+i}{t+p+(k-2)(s+1)+i}$$
$$>\frac{(s+1)2^{k-1}}{t+p+(k-2)(s+1)}(s(s+1))^s$$
$$>(s+1)2^{k-1}e^{-1}(t+p+(k-1)(s+1))^{-1}. \tag{12}$$

Elementary counting yields that for $s+1\geqslant 6c$ this ratio is greater than 1 while for $s+1\geqslant 12c$ it is at least 2. Hence

$$|\mathcal{F}|2^n<[12c]^*\max_{s=0}^{[6c]}\binom{t+p+(k-1)s}{s}2^{-(t+p+(k-1)s)}$$
$$<[12c]^*2^{-t-p}\max_{s=0}^{[6c]}((t+p+(k-1)s/2^{k-1})^s/s!). \tag{13}$$

Now using $s!>(s/e)^s$ and $(p+(k-1)s)s\leqslant(p+(k-1)6t\cdot 2^{-k})6t2^{-k}<t$ we obtain:

$$|\mathcal{F}|\cdot 2^{-n}<[12c]^*2^{-t-p}\max_{s=0}^{[6c]}((te)2^{-(k-1)}(1+(p+(k-1)s)/t)/s)^s$$
$$<[12c]^*e2^{-t-p}\max_{s=0}^{[6c]}(te2^{-(k-1)}/s)^s \tag{14}$$

The function $f(s)=(q/s)^s$ attains its maximum at $s=q/e$ whence

$$|\mathcal{F}|\cdot 2^{-n}<[12c]^*e2^{-t-p}e^{2c}=$$
$$=2^{-t-[\log([12c]^*e)+2c\log e]^*}[12c]^*e\cdot e^{2c}=2^{-t},$$

contradicting the maximality of $|\mathcal{F}|$.

5. The proof of the main theorem

Let us choose $k=1$ sets $F_1,\ldots,F_{k-1}\in\mathcal{F}$ such that their union, D is of maximal cardinality, say $n-t-h$. According to Proposition 3 we may assume $D=[1,n-t-h]$ and in view of the lemma

$$0\leqslant h\leqslant 2c\log e+\log([12c]^*e). \tag{15}$$

The property $P(n,k,t)$ implies that for any $F\in\mathcal{F}$

$$|F\cap[n-t-h+1,n]|\leqslant h. \tag{16}$$

Let q denote the greatest integer such that there exists a set $F^q\in\mathcal{F}$, $|F^q\cap[n-t-h+1,n]|=q$. $F\cap[n-t+1,n]\neq\emptyset$. We may suppose that such a q exists as otherwise $F\subseteq[1,n-t]$ holds for every $F\in\mathcal{F}$ and by the maximality of $\mathcal{F}$ it follows that $\mathcal{F}$ is $E(n,k,t,0)$. Let $0\leqslant r\leqslant q$ and let $A\subseteq[n-t-h+1,n]$, $|A|=r$. Let us set $\mathcal{F}=\{F-A\mid F\in\mathcal{F}, F\cap[n-t-h+1,n]=A\}$.

Proposition 5. *$\mathscr{E}_A$ is a family of subsets of $[1, n-t-h]$ which has property*

$$P\left(n-t-h, \left[\frac{k}{2}\right], \left[\frac{k+1}{2}\right]+kr-h\right).$$

Proof. If this statement is not true, then in view of Proposition 3 there exists sets $E_1, \ldots, E_{[k/2]} \in \mathscr{E}_A$ such that

$$\bigcup_{i=1}^{[k/2]} E_i = \left[1, n-t-kr-\left[\frac{k+1}{2}\right]+1\right]. \tag{17}$$

Let us introduce the notation

$$n-t-kr-\left[\frac{k+1}{2}\right]+1=m, \qquad \left[\frac{k}{2}\right]=d.$$

In view of Proposition 3 and the fact that $(E_i \cup A) \in \mathscr{F}$ for $i=1, \ldots, d$ and $F^q \in \mathscr{F}$ the following sets belong to $\mathscr{F}$ as well:

$$F_i = (E_i \cup \{m+(i-1)r+j \mid j=1, 2, \ldots, r\}) \cap [1, n], \quad i=1, \ldots, d,$$

$$F_p = \left\{m+dr+j\left[\frac{k+1}{2}\right]+p-d \mid j=0, \ldots, r\right\} \cap [1, n], \quad p=d+1, \ldots, k.$$

But $\bigcup_{i=1}^{k} F_i = [1, n-t+1]$ contradicting the property $P(n, k, t)$. In particular it follows that $m \geq 0$.

Now Proposition 5 and Theorem 1 imply

$$|\mathscr{E}_A| \leq 2^{n-t-h-(k+1)/2+kr-h)} = 2^{n-t-kr-(k+1)/2} \tag{18}$$

We may apply Theorem 1 as for $k \geq 20$

$$(2^{[k/2]-1}-k+1)(e+1)^{-1}-1 > \left[\frac{k+1}{2}\right]+(k-1)\left(\frac{2k}{150}\log e+\log\left[\frac{12k}{150}\right]^* e\right).$$

Let us define $\mathscr{F}' = \{F \in \mathscr{F} \mid F \subseteq [1, n-t]$. By the maximality of $\mathscr{F}$ we can find sets $F_1, \ldots, F_k \in \mathscr{F}$ such that $\bigcup_{i=1}^{k} F_i = n-t$. Using Proposition 3 and the definition of $\mathscr{F}'$ we conclude the existence of sets $F_1', \ldots, F_k' \in \mathscr{F}'$ such that $\bigcup_{i=1}^{k} F_i' = [1, n-t]$. On the other hand we may assume that there are no $k=1$ sets $G_1', \ldots, G_{k-1}'$ having $[1, n-t]$ for their union as in this case $P(n, k, t)$ implies again $\mathscr{F}$ is $E(n, k, t, 0)$. So the conditions of the above cited theorem of A. Brace and D.E. Daykin are fulfilled whence

$$|\mathscr{F}'| \leq \frac{k+1}{2^k} 2^{n-t}. \tag{19}$$

Let B be a subset of $[n-t-h+1, n]$ such that $|B|=q$, $B \cap [n-t+1, n] \neq \emptyset$. Let us set $\mathscr{E}_B = \{F - B \mid F \in \mathscr{F}, F \cap [n-t-h+1, n] = B\}$. If for every $F \in \mathscr{F}$ $|F \cap [n-t-qk+1, n]| \leq q$ holds, then by the maximality of $\mathscr{F}$

$$\mathscr{F} = \{F \subseteq X \mid |F \cap [n-t-qk+1, n]| \leq q\}$$

and we are done. So we may assume that there exists a set $F^0 \in \mathscr{F}$ such that $|F^0 \cap [n-t-kq+1, n]| \geq q+1$. Hence by Proposition 3 $F_k = \{n-t-kq+i \mid i = 1, \ldots, q+1\} \cap [1, n] \in \mathscr{F}$.

Proposition 6. *The family $\mathscr{E}_B$ has property $P(n-t-h, k-1, kq+1-h)$.*

Proof. If it is not true, then using Proposition 3 we may assume the existence of sets $E_1, \ldots, E_{k-1} \in \mathscr{E}_B$ such that $\bigcup_{i=1}^{k-1} E_i = [1, n-t-kq]$. Again by Proposition 3 the following sets belong to $\mathscr{F}$:

$$F_i = (\mathscr{E}_i \cup \{n-t-kq+q+1+j(k-1)+i \mid j = 0, \ldots, q-1\}) \cap [1, n], \qquad i = 1, \ldots, k-1.$$

Now $\bigcup_{i=1}^{k} F_i = [1, n-t+1]$ yields the desired contradiction, and in particular $n-t-k1+1 \geq 0$. In view of Theorem 1 Proposition 6 entails

$$|\mathscr{E}_B| \leq 2^{n-t-h-(kq+1-h)} = 2^{n-t-kq-1}. \tag{20}$$

We may apply Theorem 1 as for $k \geq 20$ we have

$$\frac{2^{k-1}-k+1}{e+1} - 1 > kq+1.$$

For $|\mathscr{F}|$ we have the following expression:

$$\begin{aligned} |\mathscr{F}| &= \sum_{A \subseteq [n-t-h+1, n]} |\{F \mid F \in \mathscr{F}, F \cap [n-t-h+1, n] = A\}| \\ &= \sum_{r=0}^{q-1} \sum_{\substack{|A|=r \\ A \subseteq [n-t-h+1, n]}} |\mathscr{E}_A| \\ &\quad + |\{F \mid F \in \mathscr{F}, F \subseteq [1, n-t], |F \cap [n-t-h+1, n]| \geq q\}| \\ &\quad + \sum_{\substack{|A|=q \\ B \subseteq [1, n-t] \\ A \subseteq [n-t-h+1, n]}} |\mathscr{E}_B|. \end{aligned} \tag{21}$$

From (21) using inequalities (18), (19), and (20) we obtain

$$\begin{aligned} |\mathscr{F}| &\leq \sum_{r=0}^{q-1} \binom{t+h}{r} 2^{n-t-kr(k+1)/2} + (k+1)2^{n-t-k} + \binom{t+h}{q} 2^{n-t-kg-1} \\ &< \sum_{r=0}^{q-1} 2^{-(k+1)/2} \binom{t+kr}{r} 2^{n-t-kr} + (k+1)2^{-k}2^{n-t} + \frac{1}{2}\binom{t+kq}{q} 2^{n-t-kq} \\ &\leq (q2^{-(k+1)/2}(k+1)2^{-k} + \tfrac{1}{2}) \max_{s=0}^{h} |E(n, k, t, s)| < \max_{s=0}^{h} |E(n, k, t, s)|. \end{aligned} \tag{22}$$

In establishing (22) we used that

$$q \leq h \leq p = 2c \log e + \log([12c]^* e)$$

and consequently for $t \leq k \cdot 2^k/150$ we have

$$q \cdot 2^{-(k+1)/2} + (k+1)2^{-k} + \tfrac{1}{2} < 1.$$

Now (22) gives the final contradiction which concludes the proof of Theorem 2.

Remark 1. The constant 150 can be considerably improved if we restrict ourselves to sufficiently large values of k.

Remark 2. It is easy to see that for $t < 2^k - k - 1$ in Theorem 2 the only optimal system is $E(n, k, t, 0)$. If $t = 2^k - k - 1$, then there are two optimal systems $E(n, k, t, 0)$ and $E(n, k, t, 1)$. In general, for any fixed positive ε and $k > k_0(\varepsilon)$, $1 \leq s \leq k/150$, $s2^k \leq t \leq (s+1-\varepsilon)2^k$ the only optimal family is $E(n, k, t, s)$.

References

[1] G. Katona, Intersection theorems for finite sets, Acta. Math. Acad. Sci. Hungar. 15(1964) 329–337.

[2] P. Erdös, Chao Ko and R. Rado, Intersection theorems for finite sets, Quart. J. Math. Oxford Ser. 12(1961) 313–318.

[3] A. Brace and D.E. Daykin, A finite set covering theorem I, Bull. Austral. Math. Soc. 5(1971) 197–202.

[4] D.J. Kleitman, Families of non-disjoint subsets, J. Combin. Theory. 1(1966) 153–155.